

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

LR

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶: C07H 21/04, 21/02, C07K 1/00, 16/00, C12N 1/14, 1/20, 5/00, 5/08, 15/00, C12P 21/06, C12Q 1/68	A1	(11) International Publication Number: WO 98/21227 (43) International Publication Date: 22 May 1998 (22.05.98)
(21) International Application Number: PCT/US97/20227 (22) International Filing Date: 12 November 1997 (12.11.97) (30) Priority Data: 08/748,591 13 November 1996 (13.11.96) US (71) Applicant: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA [US/US]; 22nd floor, 300 Lakeside Drive, Oakland, CA 94612-3550 (US). (72) Inventors: EPSTEIN, Ervin; 533A Miner Road, Orinda, CA 94563 (US). HU, Zhilan; 742 Judah Street, San Francisco, CA 94110 (US). BONIFAS, Jeanette; 252 Andover Street, San Francisco, CA 94110 (US). (74) Agent: SHERWOOD, Pamela, J.; Bozicevic & Reed LLP, Suite 200, 285 Hamilton Avenue, Palo Alto, CA 94301 (US).		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i>
(54) Title: MUTANT HUMAN HEDGEHOG GENE (57) Abstract A mutation in the human <i>sonic hedgehog</i> gene is associated with tumorigenesis. A variety of human tumors, including basal cell carcinomas, breast carcinomas, medulloblastomas, etc., have a somatic mutation that results in an amino acid substitution at position 133 [*his133 SHH], or in a mutation at position 114. Such mutated genes and fragments thereof, encoded protein, and antibodies specific for the mutated protein are useful in characterizing the phenotype of associated tumors. The mutant protein is useful in drug screening for compositions that antagonize or otherwise modulate HH activity or expression. The encoded protein is also used as a therapeutic, to modulate cell proliferation and differentiation, and treatment of pathological conditions associated with decreased hedgehog signaling.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

MUTANT HUMAN HEDGEHOG GENE

STATEMENT AS TO FEDERALLY SPONSORED RESEARCH

This invention was made with Government support under Grant no. AR39959, awarded by the National Institutes of Health. The Government has certain rights in this
5 invention.

INTRODUCTION

Life requires balance. At all levels, from populations, individuals, cells, and even down to biochemical signals within a cell, there is a balance of opposing forces. Replication is a particularly important process to keep under tight control, with cancer
10 serving as an archetype for uncontrolled cell growth. Unraveling both sides of the biochemical signals that regulate growth has provided insight into the mechanisms of tumor cell transformation.

Many genes involved in the regulation of growth and control of cellular signaling are also involved in oncogenesis. Such genes may be oncogenes, which are
15 typically upregulated in tumor cells, or tumor suppressor genes, which are down-regulated or absent in tumor cells. Malignancies can arise when a tumor suppressor is lost and/or an oncogene is inappropriately activated. Information about these regulatory signals can be used in cancer diagnosis and treatment. This knowledge can also be exploited to counteract situations where there is inadequate cell growth.

20 Segment polarity genes were originally discovered as mutations in flies that change the pattern of body segment structures. Mutations in these genes cause animals to develop changed patterns on the surfaces of body segments; the changes affecting the pattern along the head to tail axis. Among the genes in this class are *hedgehog*, which encodes a secreted protein (HH), and *patched* (PTC), which encodes a protein structurally
25 similar to transporter proteins. The *hedgehog* gene of flies has at least three vertebrate relatives: *Sonic hedgehog* (SHH); *Indian hedgehog* (IHH), and *Desert hedgehog* (DHH). Based on genetic experiments in flies, *patched* and *hedgehog* have antagonistic effects in development. PTC down regulates transcription of itself, members of the transforming growth factor β and *Wnt* gene families, and possibly other genes. Among other activities,

HH up regulates expression of *patched* and other genes that are negatively regulated by *patched*.

Recently, interest in the human *hedgehog* signaling pathway was piqued by the discovery of inherited *patched* gene mutations in patients with basal cell nevus syndrome, a rare autosomal dominant disease characterized by developmental abnormalities and frequent skin carcinomas. Somatic acquired mutations of the *patched* gene have been identified in sporadic cancers, including basal cell carcinomas, primary breast carcinomas, medulloblastomas and meningiomas. It is currently believed that *patched* acts as a tumor suppressor, and that these mutations cause a loss of function in the *patched* gene product.

10 The *hedgehog/patched* signaling pathway may therefore be a factor in tumorigenesis.

Detecting genetic alterations that lead to increased cell growth and tumorigenesis is of great interest for clinical medicine. Identifying the specific changes that lead to altered cell growth may open the door to improved diagnosis and possible treatment for associated tumors.

15 *Relevant Literature*

Discussion of the role of *hedgehog* may be found in Riddle *et al.* (1993) Cell 75:1401-1416; Echelard *et al.* (1993) Cell 75:1417-1430; Krauss *et al.* (1993) Cell 75:1431-1444 (1993); Tabata and Kornberg (1994) Cell 76:89-102; Heemskerk and DiNardo (1994) Cell 76:449-460; and Roelink *et al.* (1994) Cell 76:761-775.

20 Marigo *et al.* (1995) Genomics 28:44-51 provide gene sequences for the wild-type human *hedgehog* genes. *Hedgehog* proteins are synthesized as approximately 300 amino acid precursors. An amino terminal signal peptide is released, and the remaining protein is auto-catalytically cleaved to generate a membrane bound N-terminal portion and a soluble C-terminal portion (Porter *et al.* (1995) Nature 374:363-366). The N-terminal

25 portion, extending from approximately amino acid 26 to amino acid 256, appears to be the active fragment involved in signaling (Feitz *et al.* (1995) Current Biology 6:643-650). It is apparently bound to the membrane by a lipophilic cholesterol moiety attached to its cleaved C-terminus (Porter *et al.* (1996) Cell 86:21-34; Porter *et al.* (1996) Science 274:255-259). The crystal structure of the N-terminus suggests the presence of a potential

30 zinc hydrolase catalytic site (Hall *et al.* (1995) Nature 378:212-216).

Descriptions of *patched* gene function may be found in Hooper and Scott (1989) Cell 59:751-765; Nakano *et al.* (1989) Nature 341:508-513; and Goodrich *et al.* (1996) Genes Develop. 10:301-312. The association of *patched* mutations with basal cell nevus syndrome and sporadic tumors are described in Johnson *et al.* (1996) Science 272:1668-1671; Hahn *et al.* (1996) Cell 85:841-851; and Gailani *et al.* (1996) Nature Genetics 14:78-81.

SUMMARY OF THE INVENTION

Isolated nucleic acid compositions encoding mutated forms of human hedgehog genes are provided. The mutated form of the sonic hedgehog (SHH) protein product is associated with oncogenic transformation. The nucleic acid compositions, and antibodies specific for the mutated protein product are useful as diagnostics to identify HH-associated cancers. The nucleic acid compositions are also used to produce the encoded protein, which may be employed for functional mapping studies, as a therapeutic, and in studying associated physiological pathways. Down-regulation of the mutated gene activity *in vivo* is useful for therapeutic purposes.

DATABASE REFERENCES FOR NUCLEOTIDE AND AMINO ACID SEQUENCES

The complete mRNA sequence encoding human sonic hedgehog (SHH) has the Genbank accession number L38518. The gene is located on human chromosome 7q36, and has the genome database accession number GDB:456309. A partial sequence of human indian hedgehog (IHH) mRNA, 5' end, has the Genbank accession number L38517. The gene is located on human chromosome 2pter, and has the genome database accession number GDB:511203. A partial sequence of human desert hedgehog (hDHH) mRNA, has the Genbank Accession number U59748.

DESCRIPTION OF THE SPECIFIC EMBODIMENTS

Human *hedgehog* gene and gene product compositions are provided. The genes encode a mutant form of *hedgehog* that is associated with tumorigenesis. The subject genes and fragments thereof, encoded protein, and antibodies specific for the mutated protein are useful in characterizing the phenotype of associated sporadic tumors.

Certain human tumors, *e.g.* basal cell carcinomas, breast carcinomas, medulloblastomas, *etc.*, have been found to have an oncogenic mutation in the *sonic hedgehog* gene that results in an amino acid substitution at position 133 [**his133 SHH*]. Tumors may be typed or staged by detection of mutated sequences, antibody detection of abnormal protein products, and functional assays for altered *hedgehog* activity. The encoded protein is useful in drug screening for compositions that antagonize or otherwise modulate HH activity or expression. The encoded protein is also used as a therapeutic, to modulate cell proliferation and differentiation, and treatment of pathological conditions associated with decreased hedgehog signaling.

As used herein, the term *hedgehog* (HH) genes and encoded proteins shall be used to generally designate any of the mammalian *hedgehog* genes and gene products, and unless otherwise stated will be the human homolog. Hedgehog genes may be characterized by sequence similarity to known hedgehog genes, and usually have at least about 75% sequence identity at the nucleotide, more usually at least about 85% sequence identity, and may have as much as 90% sequence identity. Methods for identifying genes by nucleic acid hybridization, or by using sequence comparison algorithms, are known in the art. Known members of the Hedgehog gene family include sonic, indian and desert hedgehog. Oncogenic mutations in HH are associated with the transformation of normal cells to tumor cells. Specific examples of oncogenic mutations are described below.

Hedgehog mutations that result in an amino acid substitution or deletion at the position corresponding to his133 in SHH, will be generally referred to as [**his HH*]. This amino acid position corresponds to position 50 of the previously referenced desert hedgehog protein (DHH), and position 39 of the previously referenced indian hedgehog protein (IHH). The alignment of amino acid sequences from the three hedgehog proteins is as follows, with the histidine residue underlined, and asterisks denoting identity between two sequences.

(DHH) SEQ ID NO:1 NALAIAMNMWPGVRLRVTEGWDEDGHHAQ

(SHH) SEQ ID NO:2 NALAISVMNQWPGVKLRVTEGWDEDGHHSE

* *****

(IHH) SEQ ID NO:3 NSLAISVMNQWPGVKLRVTEGWDEDGHHSE

Specific mutations may be designated according to the affected hedgehog gene product, *i.e.* [*his SHH], [*his DHH] and [*his IHH], by the specific amino acid substitution, *i.e.* [his→tyr HH]; [Δ his HH], *etc.*, or both, *i.e.* [his→tyr SHH]. Mutations may also be specified in terms of the DNA sequence. Proteins of particular interest have a substitution of tyrosine, phenylalanine or tryptophan for [*his HH]. The substitution of other non-polar or non-charged amino acids at this position are also of interest, *i.e.* alanine, valine, leucine, isoleucine, proline, methionine, glycine, serine, cysteine, asparagine and glutamine.

The mutated form of *sonic hedgehog*, [his→tyr SHH; SEQ ID NO:4], is associated with oncogenic transformation of human cells. This is demonstrated by the presence of the mutated DNA sequence (C→T transition at nt. 548; SEQ ID NO:5) in a number of independently arising tumors. The spontaneously occurring nucleotide change encodes a tyrosine residue, but other codons can be engineered into the gene sequence, through conventional recombinant DNA technology, to provide proteins having a similar phenotype. The high degree of homology between HH proteins in this region indicates a shared structure and function, with a similar phenotypic change resulting from this amino acid change in IHH and DHH. It is believed that this mutation alters the zinc hydrolase activity of the protein.

Other mutations of interest include an amino acid substitution of the methionine found at residue 114 of SHH, herein referred to as a [*met HH] mutation, sequences shown below, with the methionine residue underlined, and asterisks denoting identity between two sequences. The known HH proteins have a high degree of sequence identity in this region. Of particular interest is a change from met→ile, *e.g.* [met→ile SHH]. The gene, SEQ ID NO:10, encodes this mutated protein, SEQ ID NO:9, and was found in the germline of a basal cell carcinoma patient.

(DHH) SEQ ID NO:6 RVNALAIAVMNMWPGVRLRVTEGWDED

***** **

(SHH) SEQ ID NO:7 KLNALAISVMNQWPGVKLRVTEGWDED

(IHH) SEQ ID NO:8 RLNSLAISVMNQWPGVKLRVTEGWDED

Nucleic Acid Compositions

The human [his→tyr SHH] amino acid sequence is provided (SEQ ID NO:4), and the encoding gene, isolated from human tumors, as (SEQ ID NO:5). In order to identify the subject oncogenic HH mutations, exonic primers from the published sequence
5 data were used to isolate *SHH* genomic clones. Sequence data from the genomic clones was used to generate specific primers. These primers were used to amplify genomic DNA from primary tumors, including breast carcinomas, basal cell carcinomas and medulloblastomas. The PCR products were screened for mutations using single strand conformation polymorphism (SSCP) analysis. The specific mutation found in SEQ ID
10 NO:5 was isolated from four independent tumors.

DNA encoding a [*his HH] protein may be cDNA or genomic DNA or a fragment thereof that encompasses the altered [*his] residue. As known in the art, cDNA sequences have the arrangement of exons found in processed mRNA, forming a continuous open reading frame, while genomic sequences may have introns interrupting the open
15 reading frame. The term "[*his HH] gene" shall be intended to mean the open reading frame encoding such specific HH polypeptides, as well as adjacent 5' and 3' non-coding nucleotide sequences involved in the regulation of expression, up to about 1 kb beyond the coding region, in either direction.

Genomic sequences of interest comprise the nucleic acids present between the
20 initiation codon and the stop codon, including all of the introns that are normally present in a native chromosome. It may include the 3' and 5' untranslated regions found in the mature mRNA. It may further include specific transcriptional and translational regulatory sequences, such as promoters, enhancers, *etc.*, including about 1 kb of flanking genomic DNA at either the 5' or 3' end of the coding region. The genomic DNA may be isolated
25 as a fragment of 50 kbp or smaller; and substantially free of flanking chromosomal sequence.

The nucleic acid compositions of the subject invention encode all or a part of the subject polypeptides. Fragments may be obtained of the DNA sequence by chemically synthesizing oligonucleotides in accordance with conventional methods, by restriction
30 enzyme digestion, by PCR amplification, *etc.* For the most part, DNA fragments will be at least about 25 nt in length, usually at least about 30 nt, more usually at least about 50 nt. For use in amplification reactions, such as PCR, a pair of primers will be used. The exact

composition of the primer sequences is not critical to the invention, but for most applications the primers will hybridize to the subject sequence under stringent conditions, as known in the art. It is preferable to chose a pair of primers that will generate an amplification product of at least about 50 nt, preferably at least about 100 nt. Algorithms
5 for the selection of primer sequences are generally known, and are available in commercial software packages. Amplification primers hybridize to complementary strands of DNA, and will prime towards each other.

The subject [*his HH] genes are isolated and obtained in substantial purity, generally as other than an intact mammalian chromosome. Usually, the DNA will be
10 obtained substantially free of other nucleic acid sequences that do not include a *hh* sequence or fragment thereof, generally being at least about 50%, usually at least about 90% pure and are typically "recombinant", *i.e.* flanked by one or more nucleotides with which it is not normally associated on a naturally occurring chromosome.

The subject nucleic acids may be used to identify expression of the gene in a
15 biological specimen. The manner in which one probes cells for the presence of particular nucleotide sequences, as genomic DNA or RNA, is well-established in the literature and does not require elaboration here. Conveniently, a biological specimen is used as a source of mRNA. The mRNA may be amplified by RT-PCR, using reverse transcriptase to form a complementary DNA strand, followed by polymerase chain reaction amplification using
20 primers specific for the subject DNA sequences. Alternatively, the mRNA sample is separated by gel electrophoresis, transferred to a suitable support, *e.g.* nitrocellulose and then probed with a fragment of the subject DNA as a probe. Other techniques may also find use. Detection of mRNA having the subject sequence is indicative of *hedgehog* gene expression in the sample.

25

Synthesis of HH Protein

The subject nucleic acid compositions may be employed for producing [*his HH] protein, or fragments thereof that encompass a [*his] mutation. For expression, an expression cassette may be employed, providing for a transcriptional and translational initiation region, which may be inducible or constitutive, the coding region under the
30 transcriptional control of the transcriptional initiation region, and a transcriptional and

translational termination region. Various transcriptional initiation regions may be employed which are functional in the expression host.

The peptide may be expressed in prokaryotes or eukaryotes in accordance with conventional ways, depending upon the purpose for expression. For large scale production
5 of the protein, a unicellular organism or cells of a higher organism, *e.g.* eukaryotes such as vertebrates, particularly mammals, may be used as the expression host, such as *E. coli*, *B. subtilis*, *S. cerevisiae*, and the like. In many situations, it may be desirable to express the subject *hedgehog* gene in a mammalian host, whereby the *hedgehog* gene product will be glycosylated, and secreted.

10 With the availability of the protein in large amounts by employing an expression host, the protein may be isolated and purified in accordance with conventional ways. A lysate may be prepared of the expression host and the lysate purified using HPLC, exclusion chromatography, gel electrophoresis, affinity chromatography, or other purification technique. The purified protein will generally be at least about 80% pure,
15 preferably at least about 90% pure, and may be up to and including 100% pure. By pure is intended free of other proteins, as well as of cellular debris.

The polypeptide is used for the production of antibodies, where short fragments provide for antibodies specific for the particular polypeptide, and larger fragments allow for the production of antibodies over the surface of the protein.
20 Antibodies may be raised to the normal or oncogenic forms. Of particular interest are antibodies that specifically recognize the mutated form of the protein, *i.e.* do not bind to the normal form. Antibodies may be raised to isolated peptides corresponding to these mutations, or to the native protein, *e.g.* by immunization with cells expressing HH, immunization with liposomes containing HH, *etc.* Such antibodies are useful in therapy
25 and diagnosis, *e.g.* typing and staging of human carcinomas.

Antibodies are prepared in accordance with conventional ways, where the expressed polypeptide or protein is used as an immunogen, by itself or conjugated to known immunogenic carriers, *e.g.* KLH, pre-S HBsAg, other viral or eukaryotic proteins, or the like. Various adjuvants may be employed, with a series of injections, as
30 appropriate. For monoclonal antibodies, after one or more booster injections, the spleen is isolated, the splenocytes immortalized, and then screened for high affinity antibody

binding. The immortalized cells, *e.g.* hybridomas, producing the desired antibodies may then be expanded. For further description, see *Monoclonal Antibodies: A Laboratory Manual*, Harlow and Lane eds., Cold Spring Harbor Laboratories, Cold Spring Harbor, New York, 1988. If desired, the mRNA encoding the heavy and light chains may be
5 isolated and mutagenized by cloning in *E. coli*, and the heavy and light chains mixed to further enhance the affinity of the antibody. Alternatives to *in vivo* immunization as a method of raising antibodies include binding to phage "display" libraries, usually in conjunction with *in vitro* affinity maturation.

Detection of HH Associated Tumors

10 DNA from a patient having a tumor suspected of association with aberrant HH signaling is analyzed for the presence of an oncogenic mutation in a *hh* gene, *e.g.* a [*his HH] mutation. Genetic characterization of sporadic tumors generally requires analysis of tumor cell DNA or RNA, conveniently with a biopsy sample. The nucleic acids are screened for the presence of an oncogenic mutation, as compared to a normal sequence
15 (previously referenced). Sporadic tumors associated with altered HH signaling include basal cell carcinomas; melanomas; squamous cell carcinomas; breast carcinomas; transitional bladder cell carcinoma; meningiomas; medullomas; fibromas of the heart and ovary; carcinomas of the lung, colon, ovary, kidney and esophagus; other carcinomas of the gut; *etc.*

20 A number of methods are available for analyzing genomic DNA sequences. Where large amounts of DNA are available, the genomic DNA is used directly. Alternatively, the region of interest is cloned into a suitable vector and grown in sufficient quantity for analysis, or amplified by conventional techniques, such as the polymerase chain reaction (PCR). The use of the polymerase chain reaction is described in Saiki, *et*
25 *al.* (1985) *Science* **239**:487, and a review of current techniques may be found in Sambrook, *et al.* *Molecular Cloning: A Laboratory Manual*, CSH Press 1989, pp.14.2-14.33.

A detectable label may be included in the amplification reaction. Suitable labels include fluorochromes, *e.g.* fluorescein isothiocyanate (FITC), rhodamine, Texas
30 Red, phycoerythrin, allophycocyanin, 6-carboxyfluorescein (6-FAM), 2',7'-dimethoxy-

4',5'-dichloro-6-carboxyfluorescein (JOE), 6-carboxy-X-rhodamine (ROX), 6-carboxy-2',4',7',4,7-hexachlorofluorescein (HEX), 5-carboxyfluorescein (5-FAM) or N,N,N',N'-tetramethyl-6-carboxyrhodamine (TAMRA), radioactive labels, *e.g.* ^{32}P , ^{35}S , ^3H ; *etc.* The label may be a two stage system, where the amplified DNA is conjugated to biotin, 5 haptens, *etc.* having a high affinity binding partner, *e.g.* avidin, specific antibodies, *etc.*, where the binding partner is conjugated to a detectable label. The label may be conjugated to one or both of the primers. Alternatively, the pool of nucleotides used in the amplification is labeled, so as to incorporate the label into the amplification product.

The amplified or cloned fragment may be sequenced by dideoxy or other 10 methods, and the sequence of bases compared to the normal *HH* sequence. Hybridization with the variant sequence may also be used to determine its presence, by Southern blots, dot blots, *etc.* Single strand conformational polymorphism (SSCP) analysis, denaturing gradient gel electrophoresis (DGGE), and heteroduplex analysis in gel matrices are used to detect conformational changes created by DNA sequence variation as alterations in 15 electrophoretic mobility. The hybridization pattern of a control and variant sequence to an array of oligonucleotide probes immobilised on a solid support, as described in WO 95/11995, may also be used as a means of detecting the presence of variant sequences. Alternatively, where an oncogenic mutation creates or destroys a recognition site for a restriction endonuclease, the fragment is digested with that endonuclease, and the product 20 size fractionated to determine whether the fragment was digested. Fractionation is performed by gel electrophoresis, particularly on acrylamide or agarose gels.

Analysis of tumor cells for the presence of aberrant *HH* proteins may be performed by immunoassay. A sample is taken from a patient suspected of having a *HH*-associated tumor. Samples, as used herein, include biological fluids such as blood, 25 cerebrospinal fluid, tears, saliva, lymph, dialysis fluid and the like; organ or tissue culture derived fluids; and fluids extracted from physiological tissues. Also included in the term are derivatives and fractions of such fluids. Biopsy samples are of particular interest, *e.g.* skin lesions, organ tissue fragments, *etc.* Where metastasis is suspected, blood samples may be preferred. The number of cells in a sample will generally be at least about 10^3 , 30 usually at least 10^4 , more usually at least about 10^5 . The cells may be dissociated, in the case of solid tissues, or tissue sections may be analyzed. Alternatively a lysate of the cells may be prepared.

Diagnosis may be performed by a number of methods. The different methods all determine the presence of abnormal HH in patient cells suspected of having an oncogenic mutation. For example, detection may utilize staining of intact cells or histological sections, performed in accordance with conventional methods. The antibodies of interest are added to the cell sample, and incubated for a period of time sufficient to allow binding to the epitope, usually at least about 10 minutes. The antibody may be labeled with radioisotopes, enzymes, fluorescers, chemiluminescers, or other labels for direct detection. Alternatively, a second stage antibody or reagent is used to amplify the signal. Such reagents are well-known in the art. For example, the primary antibody may be conjugated to biotin, with horseradish peroxidase-conjugated avidin added as a second stage reagent. Final detection uses a substrate that undergoes a color change in the presence of the peroxidase. The absence or presence of antibody binding may be determined by various methods, including flow cytometry of dissociated cells, microscopy, radiography, scintillation counting, *etc.*

An alternative method for diagnosis depends on the *in vitro* detection of binding between antibodies and HH in a cell lysate, supernatant or other fluid into which the tumor cells may secrete active HH protein. Measuring the concentration of HH binding in a sample or fraction thereof may be accomplished by a variety of specific assays. A conventional sandwich type assay may be used. For example, a sandwich assay may first attach HH-specific antibodies to an insoluble surface or support. The particular manner of binding is not crucial so long as it is compatible with the reagents and overall methods of the invention. They may be bound to the plates covalently or non-covalently, preferably non-covalently.

The insoluble supports may be any compositions to which polypeptides can be bound, which is readily separated from soluble material, and which is otherwise compatible with the overall method. The surface of such supports may be solid or porous and of any convenient shape. Examples of suitable insoluble supports to which the receptor is bound include beads, *e.g.* magnetic beads, membranes and microtiter plates. These are typically made of glass, plastic (*e.g.* polystyrene), polysaccharides, nylon or nitrocellulose. Microtiter plates are especially convenient because a large number of assays can be carried out simultaneously, using small amounts of reagents and samples.

Patient sample lysates or supernatants are then added to separately assayable supports (for example, separate wells of a microtiter plate) containing antibodies. Preferably, a series of standards, containing known concentrations of normal and/or abnormal HH is assayed in parallel with the samples or aliquots thereof to serve as

5 controls. Preferably, each sample and standard will be added to multiple wells so that mean values can be obtained for each. The incubation time should be sufficient for binding, generally, from about 0.1 to 3 hr is sufficient. After incubation, the insoluble support is generally washed of non-bound components. Generally, a dilute non-ionic detergent medium at an appropriate pH, generally 7-8, is used as a wash medium. From

10 one to six washes may be employed, with sufficient volume to thoroughly wash non-specifically bound proteins present in the sample.

After washing, a solution containing a second, detecting antibody is applied. The antibody will bind the immobilized HH with sufficient specificity such that it can be distinguished from other components present. The second antibodies may be labeled to

15 facilitate direct, or indirect quantification of binding. Examples of labels that permit direct measurement of second receptor binding include radiolabels, such as ^3H or ^{125}I , fluorescers, dyes, beads, chemiluminescers, colloidal particles, and the like. Examples of labels which permit indirect measurement of binding include enzymes where the substrate may provide for a colored or fluorescent product. In a preferred embodiment,

20 the antibodies are labeled with a covalently bound enzyme capable of providing a detectable product signal after addition of suitable substrate. Examples of suitable enzymes for use in conjugates include horseradish peroxidase, alkaline phosphatase, malate dehydrogenase and the like. Where not commercially available, such antibody-enzyme conjugates are readily produced by techniques known to those skilled in the art. The

25 incubation time should be sufficient for the labeled ligand to bind available molecules. Generally, from about 0.1 to 3 hr is sufficient, usually 1 hr sufficing.

After the second binding step, the insoluble support is again washed free of non-specifically bound material. The signal produced by the bound conjugate is detected by conventional means. Where an enzyme conjugate is used, an appropriate enzyme

30 substrate is provided so a detectable product is formed.

Other immunoassays are known in the art and may find use as diagnostics. Ouchterlony plates provide a simple determination of antibody binding. Western blots

may be performed on protein gels or protein spots on filters, using a detection system specific for HH as desired, conveniently using a labeling method as described for the sandwich assay.

Down-Regulation of HH Expression

5 Antisense molecules specific for the oncogenic HH genes, *e.g.* [*his HH], [*met HH], are used to down-regulate expression in cells suspected or shown to have [*his HH] associated tumors. Administration of the antisense molecules has the effect of decreasing the oncogenic HH activity. The anti-sense reagent may be antisense oligonucleotides (ODN), particularly synthetic ODN having chemical modifications from
10 native nucleic acids, or nucleic acid constructs that express such anti-sense molecules as RNA. The antisense sequence is complementary to the mRNA of the targeted HH gene, and inhibits expression of the targeted gene products. Antisense molecules inhibit gene expression through various mechanisms, *e.g.* by reducing the amount of mRNA available for translation, through activation of RNase H, or steric hindrance. One or a combination
15 of antisense molecules may be administered, where a combination may comprise multiple different sequences.

 Antisense molecules may be produced by expression of all or a part of the target gene sequence in an appropriate vector, where the vector is introduced and expressed in the targeted tumor cells. The transcriptional initiation will be oriented such
20 that the antisense strand is produced as an RNA molecule. The anti-sense RNA hybridizes with the endogenous sense strand mRNA, thereby blocking expression of the targeted gene. The native transcriptional initiation region, or an exogenous transcriptional initiation region may be employed. The promoter may be introduced by recombinant methods *in vitro*, or as the result of homologous integration of the sequence into a
25 chromosome. Many strong promoters are known in the art, including the β -actin promoter, SV40 early and late promoters, human cytomegalovirus promoter, retroviral LTRs, methallothionein responsive element (MRE) and tetracycline-inducible promoter constructs, *etc.*

 Transcription vectors generally have convenient restriction sites located near
30 the promoter sequence to provide for the insertion of nucleic acid sequences.

Transcription cassettes may be prepared comprising a transcription initiation region, the

target gene or fragment thereof, and a transcriptional termination region. The transcription cassettes may be introduced into a variety of vectors, *e.g.* plasmid; retrovirus, *e.g.* lentivirus; adenovirus; and the like, where the vectors are able to transiently or stably be maintained in the tumor cells, usually for a period of at least about
5 one day, more usually for a period of at least about several days to several weeks.

Alternatively, the antisense molecule is a synthetic oligonucleotide. Antisense oligonucleotides will generally be at least about 7, usually at least about 12, more usually at least about 20 nucleotides in length, and not more than about 500, usually not more than about 50, more usually not more than about 35 nucleotides in length, where the length is
10 governed by efficiency of inhibition, specificity, including absence of cross-reactivity, and the like. It has been found that short oligonucleotides, of from 7 to 8 bases in length, can be strong and selective inhibitors of gene expression (see Wagner *et al.* (1996) Nature Biotechnology 14:840-844).

A specific region or regions of the endogenous sense strand mRNA sequence
15 is chosen to be complemented by the antisense sequence, preferably encompassing the [*his HH] mutation. Selection of a specific sequence for the oligonucleotide may use an empirical method, where several candidate sequences are assayed for inhibition of expression of the target gene in an *in vitro* or animal model. A combination of sequences may also be used, where several regions of the mRNA sequence are selected for antisense
20 complementation.

Antisense oligonucleotides may be chemically synthesized by methods known in the art (see Wagner *et al.* (1993) *supra.* and Milligan *et al.*, *supra.*) Preferred oligonucleotides are chemically modified from the native phosphodiester structure, in order to increase their intracellular stability and binding affinity. A number of such
25 modifications have been described in the literature, which alter the chemistry of the backbone, sugars or heterocyclic bases.

Among useful changes in the backbone chemistry are phosphorothioates; phosphorodithioates, where both of the non-bridging oxygens are substituted with sulfur; phosphoroamidites; alkyl phosphotriesters and boranophosphates. Achiral phosphate
30 derivatives include 3'-O'-5'-S-phosphorothioate, 3'-S-5'-O-phosphorothioate, 3'-CH₂-5'-O-phosphonate and 3'-NH-5'-O-phosphoroamidate. Peptide nucleic acids replace the entire ribose phosphodiester backbone with a peptide linkage.

Sugar modifications are also used to enhance stability and affinity. The α -anomer of deoxyribose may be used, where the base is inverted with respect to the natural β -anomer. The 2'-OH of the ribose sugar may be altered to form 2'-O-methyl or 2'-O-allyl sugars, which provides resistance to degradation without comprising affinity.

5 Modification of the heterocyclic bases must maintain proper base pairing. Some useful substitutions include deoxyuridine for deoxythymidine; 5-methyl-2'-deoxycytidine and 5-bromo-2'-deoxycytidine for deoxycytidine. 5-propynyl-2'-deoxyuridine and 5-propynyl-2'-deoxycytidine have been shown to increase affinity and biological activity when substituted for deoxythymidine and deoxycytidine, respectively.

10 Oligonucleotides may additionally comprise a targeting moiety that enhances uptake of the molecule by tumor cells. The targeting moiety is a specific binding molecule, *e.g.* an antibody or fragment thereof that recognizes molecules present on the surface of tumor cells. Bispecific antibodies, chimeric antibodies and single chain antibodies are known in the art. Suitably prepared non-human antibodies can be
15 humanized in various ways. Linkage between the oligonucleotide and targeting moiety may use any conventional method, for example by disulfide, amide or thioether bonds, depending on the chemistry of the oligonucleotide backbone. Preferably the linkage will be cleaved inside the cell to liberate the oligonucleotide.

 Oligonucleotides can be conjugated to hydrophobic residues, *e.g.* cholesterol,
20 to protect from nucleases and to improve transport across cell membranes. Alternatively, conjugation to poly-L-lysine or other polyamines may also enhance delivery to the cell. A further modification that can be made is the addition of an intercalating component, such as acridine, capable of intercalating into the target mRNA and stabilizing the resultant hybrid. Antisense oligonucleotides may be transfected in combination with an enzyme(s)
25 that will degrade antisense-mRNA complexes in the cell, *e.g.* RNase H. Any protein or enzyme that can preferentially degrade or sequester the antisense-mRNA duplex may be similarly useful.

 As an alternative to anti-sense inhibitors, catalytic nucleic acid compounds, *e.g.* ribozymes, anti-sense conjugates, *etc.* may be used to inhibit gene expression.

30 Ribozymes may be synthesized *in vitro* and administered to the patient, or may be encoded on an expression vector, from which the ribozyme is synthesized in the targeted cell (for example, see International Patent Application WO 95/23225, and Beigelman et al. (1995)

Nucl. Acids Res **23**:4434-42). Examples of oligonucleotides with catalytic activity are described in WO 9506764. Conjugates of anti-sense ODN with a metal complex, *e.g.* terpyridylCu(II), capable of mediating mRNA hydrolysis are described in Bashkin *et al.* (1995) Appl Biochem Biotechnol **54**:43-56

5 The antisense molecules and/or other inhibitory agents are administered by contact with the tumor cells under conditions that permit entry. The molecules may be provided in solution or in any other pharmacologically suitable form for administration, such as a liposome suspension. There are many delivery methods known in the art for enhancing the uptake of nucleic acids by cells. Useful delivery systems include Sendai
10 virus-liposome delivery systems (see Rapaport and Shai (1994) J. Biol. Chem. **269**:15124-15131), cationic liposomes, polymeric delivery gels or matrices, porous balloon catheters (as disclosed by Shi *et al.* (1994) Circulation **90**:955-951; and Shi *et al.* (1994) Gene Therapy **1**:408-414), retrovirus expression vectors, and the like.

 The use of liposomes as a delivery vehicle is one method of interest. The
15 liposomes fuse with the cells of the target site and deliver the contents of the lumen intracellularly. The liposomes are maintained in contact with the cells for sufficient time for fusion, using various means to maintain contact, such as isolation, binding agents, and the like. Liposomes may be prepared with purified proteins or peptides that mediate fusion of membranes, such as Sendai virus or influenza virus, etc. The lipids may be any
20 useful combination of known liposome forming lipids, including cationic lipids, such as phosphatidylcholine. The remaining lipid will normally be neutral lipids, such as cholesterol, phosphatidyl serine, phosphatidyl glycerol, and the like.

 For preparing the liposomes, the procedure described by Kato *et al.* (1991) J. Biol. Chem. **266**:3361 may be used. Briefly, the lipids and lumen composition containing
25 the nucleic acids are combined in an appropriate aqueous medium, conveniently a saline medium where the total solids will be in the range of about 1-10 weight percent. After intense agitation for short periods of time, from about 5-60 sec., the tube is placed in a warm water bath, from about 25-40°C and this cycle repeated from about 5-10 times. The composition is then sonicated for a convenient period of time, generally from about
30 1-10 sec. and may be further agitated by vortexing. The volume is then expanded by adding aqueous medium, generally increasing the volume by about from 1-2 fold, followed

by shaking and cooling. This method allows for the incorporation into the lumen of high molecular weight molecules, particularly nucleic acids of one kbp or more. In this way oligonucleotides, plasmids, large genes, chromosomal fragments, viruses or viral segments may be introduced into cells efficiently.

- 5 The therapeutic agents are administered at a dose effective to reduce expression level of [*his HH] at least about 50%, more usually at least 80%, and preferably to substantially undetectable levels.

Therapeutic Use of HH Protein

The subject recombinantly produced HH protein may be therapeutically
10 administered in conditions characterized by inadequate HH signaling pathway activity, or where increased HH signaling is desirable. Examples of such uses are provided in the published International Patent Application WO95/18856, herein incorporated by reference. Hedgehog biological activity includes the ability to induce or otherwise modulate formation and differentiation of various tissues, including the head, limbs, lungs, central
15 nervous system or mesodermal patterning of embryos. Proliferation is also modulated by HH in a number of tissues. Such modulation may be achieved in *in vitro* or *in vivo* situations. For example, wound healing, bone formation, the treatment of hypoproliferative or hyperproliferative skin disorders, induction of differentiation, are affected by administration of the subject proteins.

- 20 HH is able to regulate neurogenesis, such a motor neuron inducing activity, a neuronal differentiation inducing activity or a neuronal survival promoting activity. HH also regulates organogenesis and induction of stem cell or germ cell differentiation, including the ability to induce chondrocytes or an involvement in spermatogenesis. The treatment of arthritis, *e.g.* osteoarthritis, rheumatoid arthritis, *etc.* may benefit from
25 administration of the subject proteins, and subsequent induction of chondrocytes and cartilage formation. HH is able to regulate the growth of hair by modulating the growth of cells in the hair sheath, and can be used therapeutically for this purpose.

Administration of HH induces expression of secondary signaling molecules, such as members of TGF beta family, bone morphogenetic proteins, and members of the
30 fibroblast growth factor family.

Many neurological disorders are associated with degeneration of discrete populations of neuronal elements and may be treated with the subject HH proteins. Specific disorders include traumatic injury, injury resulting from ischemia resulting from stroke, damage resulting from inflammation and/or infection of the nervous system, 5 Alzheimer's disease, Parkinson's disease, Huntington's chorea, amyotrophic lateral sclerosis, spinocerebellar degenerations, and chronic immunological diseases of the central nervous system, *e.g.* multiple sclerosis. The subject HH proteins are also useful in treating autonomic disorders of the peripheral nervous system, such as tachycardia or atrial cardiac arrhythmias arising from a degenerative condition of the nerves innervating the 10 striated muscles of the heart.

Also of interest are *in vitro* uses, where HH is added to specific cell cultures, *e.g.* neural progenitor cells, which can terminally differentiate into neurons and glia. HH sustains the reproduction of such cells in culture, in combination with appropriate culture medium, as known in the art.

15 The host may be treated with intact [*his HH] protein, or an active fragment thereof, particularly a cleaved fragment as generated by normal processing. Desirably, the peptides will not induce an immune response, particularly an antibody response. Xenogenic analogs may be screened for their ability provide a therapeutic effect without raising an immune response.

20 Various methods for administration may be employed. The polypeptide formulation may be given orally, or may be injected intravascularly, subcutaneously, peritoneally, *etc.* The dosage of the therapeutic formulation will vary widely, depending upon the nature of the disease, the frequency of administration, the manner of administration, the clearance of the agent from the host, and the like. The initial dose may 25 be larger, followed by smaller maintenance doses. The dose may be administered as infrequently as weekly or biweekly, or fractionated into smaller doses and administered daily, semi-weekly, *etc.* to maintain an effective dosage level. In many cases, oral administration will require a higher dose than if administered intravenously. The amide bonds, as well as the amino and carboxy termini, may be modified for greater stability on 30 oral administration.

The subject peptides may be prepared as formulations at a pharmacologically effective dose in pharmaceutically acceptable media, for example normal saline, PBS, *etc.*

The additives may include bactericidal agents, stabilizers, buffers, or the like. In order to enhance the half-life of the subject peptide or subject peptide conjugates, the peptides may be encapsulated, introduced into the lumen of liposomes, prepared as a colloid, or another conventional technique may be employed that provides for an extended lifetime of the

5 peptides.

The peptides may be administered as a combination therapy with other pharmacologically active agents. The additional drugs may be administered separately or in conjunction with the peptide compositions, and may be formulated in the same formulation.

10 Genetically Modified Cells and Animals

The subject nucleic acids can be used to generate transgenic animals or site specific gene modifications in cell lines. Transgenic animals may be made through homologous recombination. Alternatively, a nucleic acid construct is randomly integrated into the genome. Vectors for stable integration include plasmids, retroviruses and other
15 animal viruses, YACs, and the like. The modified cells or animals are useful in the study of *hedgehog* function and regulation. A detectable marker, such as *lac Z* may be introduced into the *hedgehog* locus, where upregulation of *hedgehog* expression will result in an easily detected change in phenotype.

DNA constructs for homologous recombination will comprise at least a
20 portion of a [*his HH] gene with the desired genetic modification, and will include regions of homology to the target locus. DNA constructs for random integration need not include regions of homology to mediate recombination. Conveniently, markers for positive and negative selection are included. Methods for generating cells having targeted gene modifications through homologous recombination are known in the art. For various
25 techniques for transfecting mammalian cells, see Keown *et al.* (1990) Methods in Enzymology 185:527-537.

For embryonic stem (ES) cells, an ES cell line may be employed, or ES cells may be obtained freshly from a host, *e.g.* mouse, rat, guinea pig, *etc.* Such cells are grown on an appropriate fibroblast-feeder layer or grown in the presence of leukemia
30 inhibiting factor (LIF). When ES cells have been transformed, they may be used to produce transgenic animals. After transformation, the cells are plated onto a feeder layer

in an appropriate medium. Cells containing the construct may be detected by employing a selective medium. After sufficient time for colonies to grow, they are picked and analyzed for the occurrence of homologous recombination or integration of the construct. Those colonies that are positive may then be used for embryo manipulation and blastocyst
5 injection. Blastocysts are obtained from 4 to 6 week old superovulated females. The ES cells are trypsinized, and the modified cells are injected into the blastocoel of the blastocyst. After injection, the blastocysts are returned to each uterine horn of pseudopregnant females. Females are then allowed to go to term and the resulting litters screened for mutant cells having the construct. By providing for a different phenotype of
10 the blastocyst and the ES cells, chimeric progeny can be readily detected.

The chimeric animals are screened for the presence of the modified gene and males and females having the modification are mated to produce homozygous progeny. The transgenic animals may be any non-human mammal, such as laboratory animals, domestic animals, etc. The transgenic animals may be used in functional studies, drug
15 screening, *etc.*, *e.g.* to determine the effect of a candidate drug on a HH-associated tumor.

Drug Screening Assays

Drug screening identifies agents that provide a replacement for HH function in abnormal cells. The role of HH as an oncogene indicates that agents which antagonize its function will inhibit the process of oncogenesis. Conversely, agents that mimic HH
20 function may stimulate controlled growth and healing. Of particular interest are screening assays for agents that have a low toxicity for human cells. A wide variety of assays may be used for this purpose, including labeled *in vitro* protein-protein binding assays, electrophoretic mobility shift assays, immunoassays for protein binding, and the like. The purified protein may also be used for determination of three-dimensional crystal structure,
25 which can be used for modeling intermolecular interactions, transporter function, *etc.*

The term "agent" as used herein describes any molecule, *e.g.* protein or pharmaceutical, with the capability of altering or mimicking the physiological function of *hedgehog*. Generally a plurality of assay mixtures are run in parallel with different agent concentrations to obtain a differential response to the various concentrations. Typically,
30 one of these concentrations serves as a negative control, *i.e.* at zero concentration or below the level of detection.

Candidate agents encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 50 and less than about 2,500 daltons. Candidate agents comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. The candidate agents often comprise cyclical carbon or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Candidate agents are also found among biomolecules including peptides, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs or combinations thereof.

Candidate agents are obtained from a wide variety of sources including libraries of synthetic or natural compounds. For example, numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides and oligopeptides. Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced. Additionally, natural or synthetically produced libraries and compounds are readily modified through conventional chemical, physical and biochemical means, and may be used to produce combinatorial libraries. Known pharmacological agents may be subjected to directed or random chemical modifications, such as acylation, alkylation, esterification, amidification, *etc.* to produce structural analogs.

Where the screening assay is a binding assay, one or more of the molecules may be joined to a label, where the label can directly or indirectly provide a detectable signal. Various labels include radioisotopes, fluorescers, chemiluminescers, enzymes, specific binding molecules, particles, *e.g.* magnetic particles, and the like. Specific binding molecules include pairs, such as biotin and streptavidin, digoxin and antidigoxin *etc.* For the specific binding members, the complementary member would normally be labeled with a molecule that provides for detection, in accordance with known procedures.

A variety of other reagents may be included in the screening assay. These include reagents like salts, neutral proteins, *e.g.* albumin, detergents, *etc.* that are used to facilitate optimal protein-protein binding and/or reduce non-specific or background interactions. Reagents that improve the efficiency of the assay, such as protease inhibitors,

nuclease inhibitors, anti-microbial agents, *etc.* may be used. The mixture of components are added in any order that provides for the requisite binding. Incubations are performed at any suitable temperature, typically between 4 and 40°C. Incubation periods are selected for optimum activity, but may also be optimized to facilitate rapid high-throughput
5 screening. Typically between 0.1 and 1 hours will be sufficient.

The compounds having the desired pharmacological activity may be administered in a physiologically acceptable carrier to a host for treatment of cancer or developmental abnormalities attributable to a defect in *hedgehog* function. The compounds may also be used to enhance *hedgehog* function in wound healing, aging, *etc.* The
10 inhibitory agents may be administered in a variety of ways, orally, topically, parenterally *e.g.* subcutaneously, intraperitoneally, by viral infection, intravascularly, *etc.* Topical treatments are of particular interest. Depending upon the manner of introduction, the compounds may be formulated in a variety of ways. The concentration of therapeutically active compound in the formulation may vary from about 0.1-100 wt. %.

15 The pharmaceutical compositions can be prepared in various forms, such as granules, tablets, pills, suppositories, capsules, suspensions, salves, lotions and the like. Pharmaceutical grade organic or inorganic carriers and/or diluents suitable for oral and topical use can be used to make up compositions containing the therapeutically-active compounds. Diluents known to the art include aqueous media, vegetable and animal oils
20 and fats. Stabilizing agents, wetting and emulsifying agents, salts for varying the osmotic pressure or buffers for securing an adequate pH value, and skin penetration enhancers can be used as auxiliary agents.

The following examples are put forth so as to provide those of ordinary skill in the art with a complete disclosure and description of how to make and use the subject
25 invention, and are not intended to limit the scope of what is regarded as the invention. Efforts have been made to insure accuracy with respect to the numbers used (*e.g.* amounts, temperature, concentrations, *etc.*) but some experimental errors and deviations should be allowed for. Unless otherwise indicated, parts are parts by weight, molecular weight is weight average molecular weight, temperature is in degrees centigrade; and pressure is at
30 or near atmospheric.

EXPERIMENTAL

Identification of Mutations in Sonic hedgehog in Human Cancers

DNA was screened from a number of human primary tumors for mutations in the *SONIC HEDGEHOG (SHH)* gene. Primers specific for the exons (coding region

5 sequences) were derived from the published gene sequence (previously referenced).

BAC clones were obtained by PCR screening of human BAC DNA pools from Research Genetics, Inc. (Huntsville, AL) (Shizuya *et al.* (1992) P.N.A.S. 89:8794-8797.

Briefly, DNA from BAC pools was used as a template in PCR with a specific exon 2 primer pair, SEQ ID NO:11; ACCGAGGGC TGGGACGAA GATGGC and SEQ ID

10 NO:12; GCGAGCCAGCATGCCGTACTTGCTG. Only pools comprising a BAC that

contains the sequence tag defined by the primer pair will yield an amplification product.

This process is continued until a single positive BAC can be identified. A positive clone, BAC270A17, was digested with restriction enzymes and ligated with vectorette linkers, as described in Riley *et al.* (1990) Nucl. Acids Res. 18:2887-2890.

15 Exon intron boundaries were determined for the three *SHH* exons by sequencing PCR products amplified using the universal vectorette primer and *SHH* cDNA primers derived from the published sequences. The exon 2-intron 3 boundary was difficult to sequence, and so a primer from the 3' end of exon 2 was used. The following primers were used to amplify genomic *SHH*:

20 exon 1; (SEQ ID NO:13) CCGCCGCGCGCACTCG
(SEQ ID NO:14) AAGGAGCGGGTGAAATCACC
exon 2; (SEQ ID NO:15) TAACGTGTCCGTCGGTGGG
(SEQ ID NO:16) TGCTTTCACCGAGCAGTGG
exon 3 (SEQ ID NO:17) CCTCCTCCCCGAGACGC
25 (SEQ ID NO:18) GGCCCCCTCCCGCGCC

These primers were used to amplify genomic DNA from primary human tumors: 34 independent basal cell carcinomas, 14 medulloblastomas and 6 breast carcinomas. PCR was performed using 1-100 ng genomic DNA, primers and AmpliTaq™ (Perkin Elmer, CA) under the following conditions: 95°C 1 min denaturation, 94°C 20
30 sec, 58°C 20 sec, 72°C 30 sec for 30 cycles followed by 72°C 2 min.

The amplification product was used in single strand conformational polymorphism (SSCP) analysis against the wild-type *SHH* gene, performed according to

the method of Chen *et al.* (1995) J. Invest Dermatol. 105:629-632. No polymorphisms were found in exons 1 and 3.

SSCP abnormalities were identified in 2 of the basal cell carcinomas, 1 of the medulloblastomas and one breast carcinoma. The aberrant SSCP bands were identical in one of the basal cell carcinomas and the two other positive tumors. No aberrantly migrating bands were found on SSCP analysis of SHH exon 2 in blood cell DNA from 100 normal persons, or from the three individuals from which the tumor cells were derived, or in exons 1 and 3. Sequencing identified the same mutation in all three tumors, a C to T transition at bp 548 (SEQ ID NO:5), predicted to cause the mutation of SEQ ID NO:4.

A second basal cell carcinoma also had an abnormally migrating band on SSCP analysis of exon 2, but the same abnormal product was also present in blood cell DNA. This is a G to A substitution at bp 493 (SEQ ID NO:9), encoding the mutated protein of SEQ ID NO:10. This individual had had at least six other basal cell carcinomas, her age at onset of the studied basal cell carcinoma was approximately 40. DNA was not available from other family members, none of whom is known to have developmental abnormalities or cancers, other than small basal cell carcinomas in both of her parents on sun exposed areas after age 60. The studied carcinoma was unusually large, involving most of one cheek, the eyelids and ear, with invasion and destruction of the facial nerve. There were no phenotypic abnormalities associated with basal cell nevus syndrome present.

SSCP analysis of the human *PATCHED* gene (*PTC*) in these four tumors identified a missense mutation in one allele of the breast carcinoma cells. Analysis of an intragenic polymorphism and direct sequencing of the mutant allele indicated a retention of both *PTC* alleles in this tumor

SHH expression in normal human epidermal keratinocytes was not detectable by Northern blot analysis under conditions in which *PTC* gene expression was easily detected, irrespective of whether the cells were cultured in low or high calcium media (0.08 or 1.2 mM). However, SHH primers did amplify a product by reverse transcriptase PCR of RNA from cultured normal human epidermal keratinocytes, and sequencing confirmed the identity of the product as authentic SHH.

DNA was extracted from keratinocytes cultured at 0.07 mM calcium using the Pharmacia Quick Prep mRNA kit. First strand cDNA was synthesized with a Gene Amp

kit (Perkin Elmer) using an oligo-dT primer. cDNA was amplified by three-stage nesting using the following primers:

Stage 1, (SEQ ID NO:19) CCGCCGCGCGCACTCG and (SEQ ID NO:20) GGCCCCCTCCCGCGCC; Stage 2, (SEQ ID NO:22) TGCCCACCGGGCAGGGGG and
5 (SEQ ID NO:21) GTCCTTCACCAGCTTGGTG; Stage 3, (SEQ ID NO:22) and (SEQ ID NO:23) TCAGCCTCCCATTGGCCGCCA.

The above results indicate that a single somatic mutation is found in cancers arising from three different tissues in independent patients. This finding indicates the importance of this mutation in carcinogenesis.

10 All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference. The citation of any publication is for its disclosure prior to the filing date and should not be construed as an admission that the present invention is not entitled to antedate such publication by virtue of
15 prior invention.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be readily apparent to those of ordinary skill in the art in light of the teachings of this invention that certain changes and modifications may be made thereto without departing from the spirit or
20 scope of the appended claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: The Regents of the University of California
- (ii) TITLE OF INVENTION: Mutant Human Hedgehog Gene
- (iii) NUMBER OF SEQUENCES: 23
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Bozicevic & Reed LLP
 - (B) STREET: 285 Hamilton Avenue, Suite 200
 - (C) CITY: Palo Alto
 - (D) STATE: CA
 - (E) COUNTRY: USA
 - (F) ZIP: 94301
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: unknown
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Sherwood, Pamela J
 - (B) REGISTRATION NUMBER: 36,677
 - (C) REFERENCE/DOCKET NUMBER: 9076/067WO1
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (415) 327-3400
 - (B) TELEFAX: (415) 327-3231

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Asn Ala Leu Ala Ile Ala Val Met Asn Met Trp Pro Gly Val Arg Leu
 1 5 10 15

Arg Val Thr Glu Gly Trp Asp Glu Asp Gly His His Ala Gln
 20 25 30

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Asn Ala Leu Ala Ile Ser Val Met Asn Gln Trp Pro Gly Val Lys Leu
 1 5 10 15

Arg Val Thr Glu Gly Trp Asp Glu Asp Gly His His Ser Glu
 20 25 30

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Asn Ser Leu Ala Ile Ser Val Met Asn Gln Trp Pro Gly Val Lys Leu
 1 5 10 15

Arg Val Thr Glu Gly Trp Asp Glu Asp Gly His His Ser Glu
 20 25 30

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 462 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

Met Leu Leu Leu Ala Arg Cys Leu Leu Leu Val Leu Val Ser Ser Leu
1           5           10           15

Leu Val Cys Ser Gly Leu Ala Cys Gly Pro Gly Arg Gly Phe Gly Lys
          20           25           30

Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys Gln Phe Ile
          35           40           45

Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Arg Tyr Glu Gly
          50           55           60

Lys Ile Ser Arg Asn Ser Glu Arg Phe Lys Glu Leu Thr Pro Asn Tyr
65           70           75           80

Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Gly Ala Asp Arg
          85           90           95

Leu Met Thr Gln Arg Cys Lys Asp Lys Leu Asn Ala Leu Ala Ile Ser
          100          105          110

Val Met Asn Gln Trp Pro Gly Val Lys Leu Arg Val Thr Glu Gly Trp
          115          120          125

Asp Glu Asp Gly Tyr His Ser Glu Glu Ser Leu His Tyr Glu Gly Arg
          130          135          140

Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Arg Ser Lys Tyr Gly Met
          145          150          155          160

Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr Glu
          165          170          175

Ser Lys Ala His Ile His Cys Ser Val Lys Ala Glu Asn Ser Val Ala
          180          185          190

Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Ala Thr Val His Leu Glu
          195          200          205

Gln Gly Gly Thr Lys Leu Val Lys Asp Leu Ser Pro Gly Asp Arg Val
          210          215          220

Leu Ala Ala Asp Asp Gln Gly Arg Leu Leu Tyr Ser Asp Phe Leu Thr
          225          230          235          240

Phe Leu Asp Arg Asp Asp Gly Ala Lys Lys Val Phe Tyr Val Ile Glu
          245          250          255

Thr Arg Glu Pro Arg Glu Arg Leu Leu Leu Thr Ala Ala His Leu Leu
          260          265          270

Phe Val Ala Pro His Asn Asp Ser Ala Thr Gly Glu Pro Glu Ala Ser
          275          280          285

```

```

Ser Gly Ser Gly Pro Pro Ser Gly Gly Ala Leu Gly Pro Arg Ala Leu
290                               295                               300

Phe Ala Ser Arg Val Arg Pro Gly Gln Arg Val Tyr Val Val Ala Glu
305                               310                               315                               320

Arg Asp Gly Asp Arg Arg Leu Leu Pro Ala Ala Val His Ser Val Thr
                               325                               330                               335

Leu Ser Glu Glu Ala Ala Gly Ala Tyr Ala Pro Leu Thr Ala Gln Gly
                               340                               345                               350

Thr Ile Leu Ile Asn Arg Val Leu Ala Ser Cys Tyr Ala Val Ile Glu
                               355                               360                               365

Glu His Ser Trp Ala His Arg Ala Phe Ala Pro Phe Arg Leu Ala His
370                               375                               380

Ala Leu Leu Ala Ala Leu Ala Pro Ala Arg Thr Asp Arg Gly Gly Asp
385                               390                               395                               400

Ser Gly Gly Gly Asp Arg Gly Gly Gly Gly Gly Arg Val Ala Leu Thr
                               405                               410                               415

Ala Pro Gly Ala Ala Asp Ala Pro Gly Ala Gly Ala Thr Ala Gly Ile
                               420                               425                               430

His Trp Tyr Ser Gln Leu Leu Tyr Gln Ile Gly Thr Trp Leu Leu Asp
                               435                               440                               445

Ser Glu Ala Leu His Pro Leu Gly Met Ala Val Lys Ser Ser
450                               455                               460

```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1576 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```

GCGAGGCAGC CAGCGAGGGA GAGAGCGAGC GGGCGAGCCG GAGCGAGGAA GGGAAAGCGC      60
AAGAGAGAGC GCACACGCAC ACACCCGCCG CGCGCACTCG CGCCCGGACC CGCACGGGGA      120
CAGCTCGGAA GTCATCAGTT CCATGGGCGA GATGCTGCTG CTGGCGAGAT GTCTGCTGCT      180
AGTCCTCGTC TCCTCGCTGC TGGTATGCTC GGGACTGGCG TGGCGACCGG GCAGGGGGTT      240
CGGGAAGAGG AGGCACCCCA AAAAGCTGAC CCCTTTAGCC TACAAGCAGT TTATCCCCAA      300

```

WO 98/21227

PCT/US97/20227

TGTGGCCGAG AAGACCCTAG GCGCCAGCGG AAGGTATGAA GGGAAGATCT CCAGAAACTC	360
CGAGCGATTT AAGGAACTCA CCCCCAATTA CAACCCCGAC ATCATATTTA AGGATGAAGA	420
AAACACCGGA GCGGACAGGC TGATGACTCA GAGGTGTAAG GACAAGTTGA ACGCTTTGGC	480
CATCTCGGTG ATGAACCAGT GGCCAGGAGT GAAACTGCGG GTGACCGAGG GCTGGGACGA	540
AGATGGCTAC CACTCAGAGG AGTCTCTGCA CTACGAGGGC CGCGCAGTGG ACATCACCAC	600
GTCTGACCGC GACCGCAGCA AGTACGGCAT GCTGGCCCCG CTGGCGGTGG AGGCCGGCTT	660
CGACTGGGTG TACTACGAGT CCAAGGCACA TATCCACTGC TCGGTGAAAG CAGAGAACTC	720
GGTGGCGGCC AAATCGGGAG GCTGCTTCCC GGGCTCGGCC ACGGTGCACC TGGAGCAGGG	780
CGGCACCAAG CTGGTGAAGG ACCTGAGCCC CGGGGACCGC GTGCTGGCGG CGGACGACCA	840
GGGCCGGCTG CTCTACAGCG ACTTCCTCAC TTTCTTGGAC CGCGACGACG GCGCCAAGAA	900
GGTCTTCTAC GTGATCGAGA CGCGGGAGCC GCGCGAGCGC CTGCTGCTCA CCGCCGCGCA	960
CCTGCTCTTT GTGGCGCCGC ACAACGACTC GGCCACCGGG GAGCCCGAGG CGTCTCGGG	1020
CTCGGGGCCG CCTTCCGGGG GCGCACTGGG GCCTCGGGCG CTGTTGCCA GCCGCGTGCG	1080
CCCCGGCCAG CGCGTGTACG TGGTGGCCGA GCGTGACGGG GACCGCCGCG TCCTGCCCCG	1140
CGCTGTGCAC AGCGTGACCC TAAGCGAGGA GGCCGCGGGC GCCTACGCGC CGCTACGGC	1200
CCAGGGCACC ATTCTCATCA ACCGGGTGCT GGCCTCGTGC TACGCGGTCA TCGAGGAGCA	1260
CAGCTGGGCG CACCGGGCCT TCGCGCCCTT CCGCCTGGCG CACGCGCTCC TGGCTGCACT	1320
GGCGCCCGCG CGCACGGACC GCGGCGGGGA CAGCGGCGGC GGGGACCGCG GGGCGGGCGG	1380
CGGCAGAGTA GCCCTAACCG CTCCAGGTGC TGCCGACGCT CCGGGTGCGG GGGCCACCGC	1440
GGGCATCCAC TGGTACTCGC AGCTGCTCTA CCAAATAGGC ACCTGGCTCC TGGACAGCGA	1500
GGCCCTGCAC CCGCTGGGCA TGGCGGTCAA GTCCAGCTGA AGCCGGGGGG CCGGGGGAGG	1560
GGCGCGGGAG GGGGCC	1576

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Arg Val Asn Ala Leu Ala Ile Ala Val Met Asn Met Trp Pro Gly Val
1 5 10 15

Arg Leu Arg Val Thr Glu Gly Trp Asp Glu Asp
20 25

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Lys Leu Asn Ala Leu Ala Ile Ser Val Met Asn Gln Trp Pro Gly Val
1 5 10 15

Lys Leu Arg Val Thr Glu Gly Trp Asp Glu Asp
20 25

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Arg Leu Asn Ser Leu Ala Ile Ser Val Met Asn Gln Trp Pro Gly Val
1 5 10 15

Lys Leu Arg Val Thr Glu Gly Trp Asp Glu Asp
20 25

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 462 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```

Met Leu Leu Leu Ala Arg Cys Leu Leu Leu Val Leu Val Ser Ser Leu
1           5           10           15

Leu Val Cys Ser Gly Leu Ala Cys Gly Pro Gly Arg Gly Phe Gly Lys
20           25           30

Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys Gln Phe Ile
35           40           45

Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Arg Tyr Glu Gly
50           55           60

Lys Ile Ser Arg Asn Ser Glu Arg Phe Lys Glu Leu Thr Pro Asn Tyr
65           70           75           80

Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Gly Ala Asp Arg
85           90           95

Leu Met Thr Gln Arg Cys Lys Asp Lys Leu Asn Ala Leu Ala Ile Ser
100          105          110

Val Met Asn Gln Trp Pro Gly Val Lys Leu Arg Val Thr Glu Gly Trp
115          120          125

Asp Glu Asp Gly His His Ser Glu Glu Ser Leu His Tyr Glu Gly Arg
130          135          140

Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Arg Ser Lys Tyr Gly Met
145          150          155          160

Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr Glu
165          170          175

Ser Lys Ala His Ile His Cys Ser Val Lys Ala Glu Asn Ser Val Ala
180          185          190

Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Ala Thr Val His Leu Glu
195          200          205

Gln Gly Gly Thr Lys Leu Val Lys Asp Leu Ser Pro Gly Asp Arg Val
210          215          220

Leu Ala Ala Asp Asp Gln Gly Arg Leu Leu Tyr Ser Asp Phe Leu Thr
225          230          235          240

Phe Leu Asp Arg Asp Asp Gly Ala Lys Lys Val Phe Tyr Val Ile Glu
245          250          255

Thr Arg Glu Pro Arg Glu Arg Leu Leu Leu Thr Ala Ala His Leu Leu
260          265          270

```

Phe Val Ala Pro His Asn Asp Ser Ala Thr Gly Glu Pro Glu Ala Ser
 275 280 285
 Ser Gly Ser Gly Pro Pro Ser Gly Gly Ala Leu Gly Pro Arg Ala Leu
 290 295 300
 Phe Ala Ser Arg Val Arg Pro Gly Gln Arg Val Tyr Val Val Ala Glu
 305 310 315 320
 Arg Asp Gly Asp Arg Arg Leu Leu Pro Ala Ala Val His Ser Val Thr
 325 330 335
 Leu Ser Glu Glu Ala Ala Gly Ala Tyr Ala Pro Leu Thr Ala Gln Gly
 340 345 350
 Thr Ile Leu Ile Asn Arg Val Leu Ala Ser Cys Tyr Ala Val Ile Glu
 355 360 365
 Glu His Ser Trp Ala His Arg Ala Phe Ala Pro Phe Arg Leu Ala His
 370 375 380
 Ala Leu Leu Ala Ala Leu Ala Pro Ala Arg Thr Asp Arg Gly Gly Asp
 385 390 395 400
 Ser Gly Gly Gly Asp Arg Gly Gly Gly Gly Gly Arg Val Ala Leu Thr
 405 410 415
 Ala Pro Gly Ala Ala Asp Ala Pro Gly Ala Gly Ala Thr Ala Gly Ile
 420 425 430
 His Trp Tyr Ser Gln Leu Leu Tyr Gln Ile Gly Thr Trp Leu Leu Asp
 435 440 445
 Ser Glu Ala Leu His Pro Leu Gly Met Ala Val Lys Ser Ser
 450 455 460

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1576 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GCGAGGCAGC CAGCGAGGGA GAGAGCGAGC GGGCGAGCCG GAGCGAGGAA GGGAAAGCGC 60
 AAGAGAGAGC GCACACGCAC ACACCCGCCG CGCGCACTCG CGCCCGGACC CGCACGGGGA 120
 CAGCTCGGAA GTCATCAGTT CCATGGGCGA GATGCTGCTG CTGGCGAGAT GTCTGCTGCT 180

WO 98/21227

PCT/US97/20227

AGTCCTCGTC TCCTCGCTGC TGGTATGCTC GGGACTGGCG TCGGGACCGG GCAGGGGGTT	240
CGGGAAGAGG AGGCACCCCA AAAAGCTGAC CCCTTTAGCC TACAAGCAGT TTATCCCCAA	300
TGTGGCCGAG AAGACCCTAG GCGCCAGCGG AAGGTATGAA GGGAAGATCT CCAGAACTC	360
CGAGCGATTT AAGGAACTCA CCCCCAATTA CAACCCCGAC ATCATATTTA AGGATGAAGA	420
AAACACCGGA GCGGACAGGC TGATGACTCA GAGGTGTAAG GACAAGTTGA ACGCTTTGGC	480
CATCTCGGTG ATAAACCAGT GGCCAGGAGT GAAACTGCGG GTGACCGAGG GCTGGGACGA	540
AGATGGCCAC CACTCAGAGG AGTCTCTGCA CTACGAGGGC CGCGCAGTGG ACATCACCAC	600
GTCTGACCGC GACCGCAGCA AGTACGGCAT GCTGGCCCCG CTGGCGGTGG AGGCCGGCTT	660
CGACTGGGTG TACTACGAGT CCAAGGCACA TATCCACTGC TCGGTGAAAG CAGAGAACTC	720
GGTGGCGGCC AAATCGGGAG GCTGCTTCCC GGGCTCGGCC ACGGTGCACC TGGAGCAGGG	780
CGGCACCAAG CTGGTGAAGG ACCTGAGCCC CGGGGACCGC GTGCTGGCGG CGGACGACCA	840
GGGCCGGCTG CTCTACAGCG ACTTCCTCAC TTTCCTGGAC CGCGACGACG GCGCCAAGAA	900
GGTCTTCTAC GTGATCGAGA CGCGGGAGCC GCGCGAGCGC CTGCTGCTCA CCGCCGCGCA	960
CCTGCTCTTT GTGGCGCCGC ACAACGACTC GGCCACCGGG GAGCCCGAGG CGTCTCGGG	1020
CTCGGGGCCG CCTTCGGGG GCGCACTGGG GCCTCGGGCG CTGTTGCGCA GCCGCGTGCG	1080
CCCGGGCCAG CGCGTGTACG TGGTGGCCGA GCGTGACGGG GACCGCCGGC TCCTGCCCCG	1140
CGCTGTGCAC AGCGTGACCC TAAGCGAGGA GGCCGCGGGC GCCTACGCGC CGCTCACGGC	1200
CCAGGGCACC ATTCTCATCA ACCGGGTGCT GGCCTCGTGC TACGCGGTCA TCGAGGAGCA	1260
CAGCTGGGCG CACCGGGCCT TCGCGCCCTT CCGCCTGGCG CACGCGCTCC TGGCTGCACT	1320
GGCGCCCGCG CGCACGGACC GCGGCGGGGA CAGCGGCGGC GGGGACCGCG GGGGCGGCGG	1380
CGGCAGACTA GCCCTAACCG CTCCAGGTGC TGCCGACGCT CCGGGTGCGG GGGCCACCGC	1440
GGGCATCCAC TGGTACTCGC AGCTGCTCTA CCAAATAGGC ACCTGGCTCC TGGACAGCGA	1500
GGCCCTGCAC CCGCTGGGCA TGGCGGTCAA GTCCAGCTGA AGCCGGGGGG CCGGGGGAGG	1560
GGCGCGGGAG GGGGCC	1576

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ACCGAGGGCT GGGACGAAGA TGGC

24

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GCGAGCCAGC ATGCCGTACT TGCTG

25

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 16 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CCGCCGCGCG CACTCG

16

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AAGGAGCGGG TGAAATCACC

20

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

TAACGTGTCC GTCGGTGGG

19

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

TGCTTTCACC GAGCAGTGG

19

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 17 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CCTCCTCCCC GAGACGC

17

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 16 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GGCCCCCTCC CGCGCC

16

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 16 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CCGCCGCGCG CACTCG

16

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 16 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GGCCCCCTCC CGCGCC

16

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

WO 98/21227

PCT/US97/20227

GTCCTTCACC AGCTTGGTG

19

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TGCCCACCGG GCAGGGGG

18

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

TCAGCCTCCC ATTTGGCCGC CA

22

WHAT IS CLAIMED IS:

1. An isolated nucleic acid encoding an oncogenic human HH protein, as other than an intact chromosome.
2. An isolated nucleic acid according to Claim 1, encoding a human [*his
5 HH] protein or a fragment of at least 18 nucleotides derived therefrom and encompassing said [*his HH] mutation, as other than an intact chromosome.
3. An isolated nucleic acid according to Claim 1, encoding a human [*met HH] protein or a fragment of at least 18 nucleotides derived therefrom and encompassing said [*met HH] mutation, as other than an intact chromosome.
- 10 4. An isolated nucleic acid according to Claim 1, wherein said protein is selected from the group consisting of sonic hedgehog, indian hedgehog and desert hedgehog.
5. An isolated nucleic acid according to Claim 1, encoding a protein selected from the group consisting of human [his→tyr HH] and [met→ile HH].
- 15 6. An isolated nucleic acid according to Claim 5, wherein said protein is selected from the group consisting of SEQ ID NO:4 and SEQ ID NO:9.
7. An isolated nucleic acid according to Claim 6, selected from the group consisting of SEQ ID NO:5 and SEQ ID NO:10.
8. An expression cassette comprising a transcriptional initiation region, a
20 nucleic acid according to Claim 1 under the transcriptional regulation of said transcriptional initiation region, and a transcriptional termination region.
9. A cell comprising an expression cassette according to Claim 8.
10. A cell according to Claim 9, wherein said cell is a mammalian cell.

11. A cell according to Claim 10, wherein said mammalian cell is human.
12. A method for producing oncogenic HH protein, said method comprising:
growing a cell according to Claim 9, whereby said oncogenic HH protein is
5 expressed; and
isolating said oncogenic HH protein free of other proteins.
13. A purified polypeptide composition comprising at least 50 weight % of the protein present as an oncogenic HH protein.
14. A purified polypeptide composition according to Claim 13, wherein
10 said protein is selected from the group consisting of sonic hedgehog, indian hedgehog, and desert hedgehog.
15. A purified polypeptide composition according to Claim 13, wherein said protein is selected from the group consisting of human [his→tyr HH] and [met→ile HH].
16. A purified polypeptide composition according to Claim 13, wherein
15 said protein is selected from the group consisting of SEQ ID NO:4 and SEQ ID NO:9.
17. A method for characterizing the phenotype of a tumor, the method comprising:
20 detecting the presence of an oncogenic HH mutation in said tumor,
wherein the presence of said oncogenic HH mutation indicates that said tumor is associated with aberrant HH signaling.
18. A method according to Claim 17, wherein said detecting step
determines the presence of said oncogenic HH mutation in a nucleic acid, selected from the
25 group consisting of genomic DNA and mRNA.

19. A method according to Claim 18, wherein said detecting step determines the presence of mutated [his→tyr SHH] protein.
20. An isolated antibody specific for a human oncogenic HH protein.
21. An isolated antibody according to Claim 18, wherein said antibody is a
5 monoclonal antibody.
22. An assay device comprising:
a support surface; and
the antibody of claim 21 bound to said support surface.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US97/20227

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :Please See Extra Sheet.

US CL :Please See Extra Sheet.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6, 69.1, 252.3, 254.11, 320.1, 325, 366; 530/350, 387.7, 387.9, 388.1; 536/23.1, 23.5; 935/22

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Databases: APS, CAS ONLINE, MEDLINE

Keywords: hedgehog, sonic, desert, imdiam, vectors, genes, DNA, deoxyribo?, human, basal cell nevus, basal cell carcinoma

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO/9518856 A1 (PRESIDENT AND FELLOWS OF HARVARD COLLEGE) 13 July 1995. see whole document.	1-22
Y, P	ORO et al. Basal cell carcinomas in mice overexpressing sonic hedgehog. Science. 02 May 1997. Vol. 276. see whole document.	1-19
A	GAILANI et al. The role of the human homologue of Drosophila patched in sporadic basal cell carcinomas. Nature Genetics. September 1996. Vol. 14. see whole document.	1-22
A, P	HAMMERSCHMIDT et al. The world according to hedgehog. Trends in Genetics. January 1997. Vol. 13. No. 1. see whole document.	1-22

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
B earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*A* document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means	
P document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

08 JANUARY 1998

Date of mailing of the international search report

18 FEB 1998

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized Officer

Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US97/20227

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A, P	STONE et al. The tumour-suppressor gene patched encodes a candidate receptor for Sonic hedgehog. Nature. 14 November 1996. Vol. 384. see whole document.	1-22
A	OGURA et al. Evidence that Shh cooperates with retinoic acid inducible co-factor to establish ZPA-like activity. Development. February 1996. Vol. 122. see whole document.	1-22
A	HAHN et al. Mutations of the human homolog of Drosophila patched in the nevoid basal cell carcinoma syndrome. Cell. 14 June 1996. Vol. 85. see whole document.	1-7
A, P	ROESSLER et al. Mutations in the human sonic hedgehog gene cause holoprosencephaly. Nature Genetics. November 1996. Vol. 14. see whole document.	1-7
X	MARIGO et al. Cloning, expression and chromosomal location of SHH and IHH: Two human homologues of the Drosophila segment polarity gene hedgehog. Genomics. 01 July 1995. Vol. 28. see whole document.	1, 4 and 8

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/20227

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
☒ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US97/20227

A. CLASSIFICATION OF SUBJECT MATTER:

IPC (6):

C07H 21/04, 21/02; C07K 1/00, 16/00; C12N 1/14, 1/20, 5/00, 5/08, 15/00; C12P 21/06; C12Q 1/68

A. CLASSIFICATION OF SUBJECT MATTER:

US CL :

435/6, 69.1, 252.3, 254.11, 320.1, 325, 366; 530/350, 387.7, 387.9, 388.1; 536/23.1, 23.5; 935/22

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claims 1-19, drawn to human hedgehog proteins, nucleic acids encoding human hedgehog proteins, vectors, transformed host cells and expression systems, thereof and a method of use.

Group II, claims 20-22, drawn to antibodies against human hedgehog proteins.

The inventions listed as Groups I and II do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: the special technical feature of the Group I invention is the specific amino acid sequence of the oncogenic mutant human hedgehog proteins and the specific nucleotide sequence of the nucleic acids encoding the human hedgehog proteins, whereas the special technical feature of the Group II invention is the antibodies against the oncogenic human hedgehog proteins. Since the special technical feature of the Group I invention is not present in the Group II claims, and the special technical feature of the Group II invention is not present in the Group I claims, unity of invention is lacking.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
223